A Network Pricing Game for Selfish Traffic *

Ara Hayrapetyan[†] Éva Tardos[‡] Tom Wexler[§]

Abstract

The success of the Internet is remarkable in light of the decentralized manner in which it is designed and operated. Unlike small scale networks, the Internet is built and controlled by a large number of disperate service providers who are not interested in any global optimization. Instead, providers simply seek to maximize their own profit by charging users for access to their service. Users themselves also behave selfishly, optimizing over price and quality of service. Game theory provides a natural framework for the study of such a situation. However, recent work in this area tends to focus on either the service providers or the network users, but not both. This paper introduces a new model for exploring the interaction of these two elements, in which network managers compete for users via prices and the quality of service provided. We study the extent to which competition between service providers hurts the overall social utility of the system.

Keywords: Game Theory, Network Pricing Games, Price of Anarchy.

* A preliminary version of this paper appeared in the Proceedings of 24th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing, July 17-20, 2005, Las Vegas, Nevada, USA.

[†]Department of Computer Science, 4126 Upson Hall, Cornell University, Ithaca, NY 14853. E-mail: ara@cs.cornell.edu. Supported in part by NSF ITR grant CCR-0325453.

[‡]Department of Computer Science, 4126 Upson Hall, Cornell University, Ithaca, NY 14853. E-mail: eva@cs.cornell.edu. Supported in part by NSF grant CCR-0311333, ITR grant CCR-0325453, and ONR grant N00014-98-1-0589.

[§]Department of Computer Science, 4126 Upson Hall, Cornell University, Ithaca, NY 14853. E-mail: wexler@cs.cornell.edu. Supported in part by NSF ITR grant CCR-0325453.

1 Introduction

One of the most surprising features of the Internet is how effectively it operates on a global scale, despite the fact that its various components (Autonomous Systems) are operated by separate service providers, each of whom seeks only to maximize their own income. A number of recent papers have considered competitive network design games as simple models of how selfish agents might construct such a network. These models tend to assume a static user population, typically with fixed demands. In this paper we address the fact that the potential network users are just as crucial to the construction of a network as the service providers themselves. Network managers compete for users via prices and the quality of service provided. We propose a simple game to explore a basic question that arises in this situation: how is the quality of service affected when the service providers set prices so as to extract maximum profit? We consider the special case of a graph with parallel links (or scheduling parallel machines), where the quality of the service is modeled as delay that increases linearly with the congestion, and user demand is concave. For such a game, we show that a pure equilibrium exists, and we provide a constant factor bound on the *price of anarchy*, which roughly speaking measures the inefficiency of competitive play. We give an improved bound for the special case in which delay is a pure congestion effect (when delay is 0 with no congestion).

Selfish routing and network design are two important classes of network games that have received much attention in recent years. In work on selfish routing (or load balancing) [10, 11, 12, 16, 19, 20, 21, 22] users in a network route their traffic selfishly with the aim of minimizing their latency. These papers show bounds on the price of anarchy in the corresponding games, that is, they give bounds on the performance degradation caused by the selfish routing as compared to a centrally designed optimal solution. In these games the sole selfish goal of users is to minimize the delay; in particular, user demand is fixed (independent of the delay in the system), and the network is passive, in that it does not try to effect routing behavior by changing the properties of its edges.

Network games have also been used to try to understand the quality of networks built by selfish agents. The price of anarchy or price of stability for such network creation games were studied under a few related models [5, 6, 13]. Such network creation games aim to model the behavior of agents, such as Autonomous Systems, who seek to build networks. However, the models considered here also assume that demand is fixed, and not dependent on the properties of the contructed network. In particular, network design games typically cannot model congestion effects. Furthermore, these models assume that network builders want to simply build a cheap network satisfying user demand, rather than maximizing their own income by charging users for access to this network.

The primary motivation of this paper is to develop a model that encompasses key aspects of both selfish routing and network design games. Our main result is a small constant bound on the price of anarchy in a game that combines profit-maximizing edge-pricing players with user demand that is sensitive to both prices and congestion. In our model, users perceive the quality of a route (path) in a network via a combination of prices and congestion. Edges have congestion sensitive delays, i.e. the time required to traverse an edge depends on the amount of traffic using it. To this extent, our model extends the work on selfish routing discussed above. However, we will also assume that each edge is operated by a distinct selfish player, who can charge traffic for the use of that edge. Finally, we assume that user demand is affected by the quality of service provided, in that increasing prices and delay lead to decreasing user traffic. We assume that the edges set prices in a selfish manner, aiming to maximize their income. The goal of our work is to quantify the performance degradation caused by selfishness (the price of anarchy) in this model. We show that in the special case of networks with parallel links and linear delays, pure Nash equilibria exist and the price of anarchy is bounded by a small constant.

Our Model.

We define a simple model that combines the issues mentioned above. We consider a network consisting of two nodes s and t, with k parallel links. Each link is controlled by a distinct player, who can charge traffic a price for use of her link. One can also view this game as modeling a type of selfish load balancing problem. Here the edges correspond to machines, and the flow or traffic will selfishly balance between them. Each machine has a load dependent delay, and it can charge a price to all users.

Link i (or machine i) has a latency (or delay) function $\ell_i(x)$, indicating the delay experienced by a volume of x traffic using i. We will primarily focus on strictly increasing linear latencies, i.e. $\ell_i(x) = a_i x + b_i$, where $a_i > 0$ and $b_i \geq 0$. We assume that user experience in routing flow through link i depends on the sum of the price and the latency, which we will call the *disutility*. More precisely, if player i charges p_i and f_i volume of flow uses link i, then that flow experiences a disutility of $p_i + \ell_i(f_i)$.

We assume that traffic routes itself *selfishly*, meaning that traffic will not route along one link if it can switch links and incur a lower disutility. As a result, all traffic will necessarily experience the same disutility.

Finally, we also assume that the total amount of traffic from s to t is dependent on the disutility that traffic experiences; as the disutility increases, the total amount of flow interested in routing from s to t decreases. We model this with a demand function $D(y)$, also referred to as the *demand curve*, which indicates the amount of flow willing to incur a disutility of y. We will naturally assume that demand $D(y)$ is decreasing in disutility y. We will focus on demand curves that are continuous and concave. Different demand curves are used to model demand in different industries. A concave demand curve is applicable for modeling demand for a service with a comparable alternative (at a high enough price all users will switch to the alternate service).

It will often be useful for us to consider $u(x) = D^{-1}(x)$, which we call the *disutility curve*. This measures the disutility that will be tolerated by a volume of x flow. By definition, $u(x)$ is also decreasing and concave.

To define our problem more precisely, we say that a price vector p induces a flow vector f satisfying the following properties.

1. For any i, j if $f_i > 0$, then $\ell_i(f_i) + p_i \leq \ell_j(f_j) + p_j$.
2. If $f_i > 0$ then $\ell_i(f_i) + p_i = u(\sum_j f_j)$.

The first condition ensures that no traffic can decrease its disutility by rerouting. The second condition states that the disutility experienced by any traffic must match the the disutility that is tolerated by the given volume of flow. It is not hard to see that since the disutility is continuous and decreasing, such flows always exist. Furthermore, Lemma 2.1 will argue that these flows are unique as well.

We may now complete the definition of the game. Each player i selects a price p_i for her link. These prices, together with the latencies and the disutility curve, induce a unique flow f_i on each link. Players seek to maximize their profit, $\pi_i = p_i \cdot f_i$. We say that a set of prices p is at *Nash equilibrium* if, by changing a single price p_i to p_i', the resulting flow f' does not give player i a larger profit ($\pi_i' = p_i' \cdot f_i'$).

For a set of prices, we are interested in measuring the *social utility* of a given solution. For a solution with prices p inducing total flow $F = \sum_j f_j$, and disutility d, we define the social utility to be

$$U(p) = \sum_i \pi_i + \int_0^F (u(x) - d)dx.$$

The first term accounts for players' profits, and the second terms represents the utility gathered by the traffic that gets routed. An example of an instance of the game with two links is shown in Figure 1. The social utility of this instance is indicated by the shaded area.

We will be interested in bounding the *price of anarchy* of this game. If p^* is the set of prices that maximizes social utility, then the *price of anarchy* is the maximum possible ratio of

$$\max_p \frac{U(p^*)}{U(p)},$$

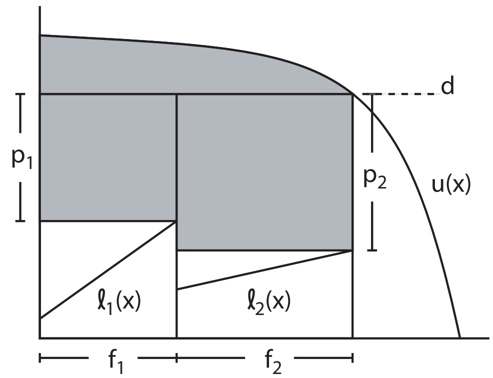

Figure 1: An instance of the network pricing game with two players.

where the prices p range over the possible equilibrium prices. Note that the total area under $u(x)$ provides an upper bound on the maximum possible social utility. However, to achieve this bound we would need to route the maximum possible traffic volume without incurring any delay.

We note that there are other reasonable measures of social ultility; we could have considered either the players' profits or the users' utility as social welfare functions. However, the sum of these two objectives seems the most reasonable. Furthermore, under either measure alone, simple examples demonstrate an unbounded price of anarchy.

We can think of our problem as a type of two-stage game where there are two types of players: the owners of the edges, and the users with traffic. In the first stage, the edge-players set prices p on edges. In the second stage, traffic routes itself selfishly from the source to the sink with respect to new latencies $\ell_i(x) + p_i$, where the rate of flow is dictated by the disutility curve. For the remainder of this paper, we will focus on the first stage of this game, the game of setting prices, and will assume that players anticipate the flow resulting from their chosen prices. However, the existence and uniqueness of these flows follow directly from observing this connection.

Our Results.

- We show that for our pricing game, in a network with parallel links, concave disutility and linear delays, there is always a pure Nash equilibrium.
- Furthermore, we bound the price of anarchy in this game by 4.65, even when delays are relaxed to be convex.
- In the special case when delay is exclusively a congestion effect, that is, all links have $\ell(0) = 0$, the price of anarchy is bounded by at most 3.125.

Related Work.

In our model traffic evaluates its experience via a combination of price and delay. Modelling traffic as experiencing disutility in terms of price and delay has been studied in the transportation literature as early as in 1920 [18] (see also [7]) and more recently in a sequence of papers started by Cole, Dodis and Roughgarden [8, 9] (see also [14, 15]). These works view prices as a tax that is set by a benevolent network manager so as to improve network performance. Acemoglu and Ozdaglar [2] consider a version of our pricing problem in a monopolistic setting with fixed demand. They assume that all edges are owned by a single service provider, and focus on establishing the existence of an equilibrium and characterizing this equilibrium. More recently and independent of this work, the same authors extended their model to oligopolies, while still assuming fixed demand and delays which are exclusively due to congestion, i.e. $\ell_i(0) = 0$ for all i (see [1]). They thereby analyzed a special case of our model and proved a tight bound of 6/5 for the price of anarchy. Along different lines, Vetta [23] shows a nice bound on the price of anarchy in a two-stage pricing game for facility location. In Vetta's game the players' strategies are the facilities they select, and prices are determined based on the facility locations, much as our prices determine the flow through the system.

There is a large body of economics literature dedicated to understanding the effects of pricing with service delays, with the focus being on establishing the existence of stable equilibria, and considering qualitative properties of equilibria (such as whether improved service leads to improved profit). Lee and Mason [17] consider the most closely related model with identical links and non-uniform users. More complex models are discussed in the recent papers by Allon and Federgruen [4] and Afèche [3] and their references. Just as in our work, these papers assume that user experience depends on a combination of price and delay, and that prices are set by selfish, income-maximizing players. However, these papers consider more complex environments, including heterogeneous users, and situations in which not only prices, but also delays can be used as strategic variables. However, unlike our work, these papers do not provide bounds on the quality degradation caused by selfish pricing.

Paper Organization.

The remainder of this paper is organized as follows. In section 2, we prove that pure Nash equilibria always exist in this game. In doing so, we first prove a number of lemmas that will be useful throughout the paper. We present our main result in section 3, where we argue that the price of anarchy in this game is bounded by a small constant. In section 4, we consider alternate classes of disutility and latency functions. In particular, we show that pure equilibria may not exist if latencies are convex, although when they do, our price of anarchy result still applies. We also argue that in the presence of convex disutility, the price of anarchy may be unbounded.

2 Existence of Pure Equilibria

We start by proving that a given price vector induces a unique flow vector.

Lemma 2.1 *For a given price vector p there is a unique flow vector f satisfying conditions 1 and 2 above.*

Proof : For a given value for disutility d, the total amount of flow F can be calculated in two ways. First, the disutility curve gives us the total flow value of $u^{-1}(d)$. Second, on each individual link i, we know that the flow has to be $max(\ell_i^{-1}(d - p_i), 0)$, and hence the total flow must be $\sum_i max(\ell_i^{-1}(d - p_i), 0)$. Since disutility is decreasing, the first function is monotone decreasing in d. Furthermore, since latencies are strictly increasing and continuous, the second function is strictly increasing in d. Hence they can have at most a single intersection point. ■

We say that a player is *content* if she has no incentive to deviate from her current price. To prove that pure Nash equilibria exist, it will be critical to relate the price charged by a content player to the flow she receives, and to the slopes of the latency and disutility functions. This is achieved through the following technical lemma, which will also play a key role in later bounding the price of anarchy.

Lemma 2.2 *Let p be the price vector chosen by the players and f be the corresponding flow vector, with total flow $F = \sum_i f_i$. Define v^- and v^+ respectively as the left and right derivatives (slopes) of the disutility curve at F. We will assume that v^- and v^+ are both well defined (note that they are both negative), though they do not have to be equal. If player i is content, then the following two conditions must hold:*

1. $(\frac{p_i}{f_i} - a_i)(\sum_{j \neq i} \frac{1}{a_j} - \frac{1}{v^-}) \geq 1$

2. $(\frac{p_i}{f_i} - a_i)(\sum_{j \neq i} \frac{1}{a_j} - \frac{1}{v^+}) \leq 1.$

Observe that equality must hold if the disutility curve is continuously differentiable at F.

Before proving this, we establish some basic results concerning the monotonicity and continuity of this game. We first show that if a single player changes her link price, the flow induced by all players' prices changes in the natural way. In particular, if a player increases her price, the newly induced flow will route less traffic on her link, at least as much traffic on every other link, and less traffic in total. Similarly, the symmetric claims hold if the player decreases her price. More precisely we have

Lemma 2.3 *Let $p_1, \ldots, p_k$ be a price vector with associated flow vector $f_1, \ldots, f_k$. Assume that the first player increases her price to $p_1' > p_1$ while the others keep their prices unchanged. Denote the new flow vector by $f_1', \ldots, f_k'$. If $f_1 > 0$, then,*

1. $f_1' < f_1$,

2. $f_i' \geq f_i$, *for all* $i \neq 1$.

3. $F' \leq F$, *where* $F = \sum_i f_i$ *and* $F' = \sum_i f_i'$.

The symmetric claim holds if player 1 reduces her price.

Proof : We will assume without loss of generality that there are at least two players with positive flow, since otherwise the claims are trivial.

1) Suppose the flow on link 1 does not decrease. Then, since the price increased, the disutility must strictly increase. Consider any other link that had positive flow. This link now has greater disutility, and since the price has not changed, must carry more flow. But this implies that both the disutility and total flow have increased, which is a contradiction.

2) Suppose some link $i \neq 1$ loses flow. Since p_i is unchanged and latencies are strictly increasing, the disutility must strictly decrease. This implies that all links that carried flow must now carry less flow. Since we know that link 1 loses flow, the total flow decreases. But then both the disutility and the total flow have decreased, which is a contradiction.

3) Finally, since flow on any link other than 1 can only increase, the disutility can only increase, and thus the total volume of flow can only decrease. ■

The next lemma argues that the amount of flow routed along any link is a continuous function of the players' prices.

Lemma 2.4 *Let p be a price vector and let $f_i(p_i)$ be the amount of flow routed through link i as a function of p_i (assuming all other prices in p are fixed). Then the function $f_i(p_i)$ is non-increasing and continuous for any p.*

Proof : The non-increasing part follows from Lemma 2.3. As for continuity, we again prove by contradiction. We will argue that if there is a discontinuity, a small change in price can be found that violates monotonicity of disutility. For some price vector p let p_i be a point of discontinuity of $f_i(\cdot)$. Let f be the flow vector corresponding to the prices p with player i charging p_i. Assume without loss of generality that $f_i(\cdot)$ is upper-discontinuous at p_i. Then there exists a small $\Delta > 0$ such that for any $\epsilon > 0$, if player i increases p_i by ϵ, the flow going through the i^{th} link will decrease by at least Δ (notice that it must decrease by Lemma 2.3). Since latencies are strictly monotone, the loss of at least Δ flow decreases the latency on

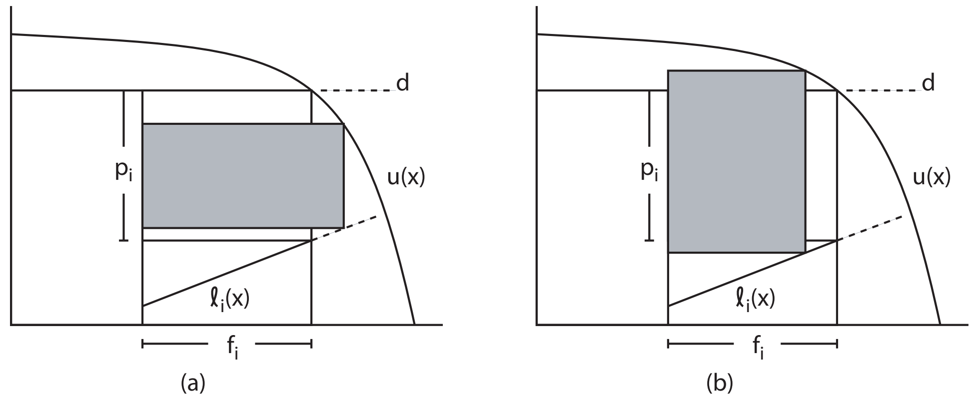

Figure 2: The immediate effect of a player (a) decreasing and (b) increasing her price.

link i by at least some $L > 0$. Choose $\epsilon < L$. Then if player i selects a price of $p_i + \epsilon$, the resulting latency will more than offset the increase in price. Thus the disutility will strictly decrease, contradicting Lemma 2.3. ■

Corollary 2.5 *The profit $\pi_i(p_i) = p_i \cdot f_i(p_i)$ of player i as a function of the price she chooses, given a fixed price vector p for the other players, is continuous for any p.*

The next lemma argues that the not only the profit of player i, but also that of any other player j changes continuously as i changes her price, while other players (including j) maintain their existing prices.

Lemma 2.6 *The profit of any other player j as a function of the price chosen by player i (assuming that the other prices are fixed), is continuous and non-decreasing.*

Proof : The non-deceasing part follows from Lemma 2.3. Let $f_i(p_i)$ and $f_j(p_i)$ be the flows of the players i and j as a function of the price p_i chosen by player i. Since $\ell_i(.)$ is continuous and player i's flow continuously depends on p_i (Lemma 2.4), $\ell_i(f_i(p_i)) + p_i$ also continuously depends on p_i. On the other hand, since both links i and j carry non-zero flow, $\ell_i(f_i(p_i)) + p_i = \ell_j(f_j(p_i)) + p_j$, where p_j is the fixed price of player j. Thus, $\ell_j(f_j(p_i))$ is also continuous in p_i. Furthermore, since $\ell_j(.)$ is strictly monotone, $f_j(p_i)$ must be continuous in p_i. Once the flow on link j is continuous, the continuity of the profit of player j is immediate. ■

Having shown that player profits depend continuously on the price of any player, we can extend this trivially to the following general profit continuity lemma.

Lemma 2.7 *Let $\pi : \mathbb{R}^k \mapsto \mathbb{R}^k$ be the function mapping a vector of prices p to the vector of profits obtained by the players when they charge the corresponding prices, i.e. $\pi(p) = (\pi_1, \ldots, \pi_k)$ means that the profit of player i is π_i under the price vector p. Then the function $\pi(\cdot)$ is a continuous function of p.*

This continuity allows us to prove Lemma 2.2.

Proof of Lemma 2.2 : The intuition behind the proof of this lemma is based on the following observations. For a player i to be content, she must not be able to benefit by changing her price at all. The lemma will follow from observing that, in particular, she will not benefit from very small changes to her price. Intuitively, the gain from an increase in price must be outweighed by the resulting loss of flow. In Figure 2 the shaded boxes indicate the immediate effect of a player increasing or decreasing her price *without* taking into account the resulting change in flow of the other players. The actual resulting flow depends not only on player i's latency function, but also on the disutility curve (determining how much flow leaves the system altogether), and other players' latencies (determining how much of the flow of player i transfers to them; flatter latency players will steal more flow than steeper latency ones). For sufficiently small increases in price, the magnitude of this loss is dictated solely by the local slopes of these curves, resulting in the first inequality. Likewise, the second inequality follows by considering a small decrease in price.

More formally, let us prove the second inequality. The proof for the first one is analogous. Assume that player i changes her price so that the system disutility is decreased by a tiny amount $\delta > 0$. Then the flow of a player $j \neq i$ will decrease by δ/a_j and the total flow in the system will increasey by $-\delta/v^+$ [1]. Hence player i's flow must increase by $\Delta f_i = \delta(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+})$. This increase in flow will result to a latency increase of $a_i \cdot \delta(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+})$. So i must decrease her price by $\Delta p_i = \delta + a_i \cdot \delta(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+})$.

We next observe that for any $\epsilon > 0$, there exists a small enough δ such that, $\Delta p_i/\Delta f_i \geq p_i/f_i - \epsilon$. Indeed, if $\Delta p_i/\Delta f_i < p_i/f_i - \epsilon$ then $p_i \Delta f_i > f_i \Delta p_i + f_i \Delta f_i \epsilon$. Now player i can obtain more profit by setting her price to $p_i - \Delta p_i$, because her new profit would be $(p_i - \Delta p_i)(f_i + \Delta f_i) = p_i f_i - f_i \Delta p_i + p_i \Delta f_i - \Delta p_i \Delta f_i > p_i \cdot f_i$ if we choose δ small enough to ensure that $\epsilon f_i > \Delta p_i$ (which we can do due to continuity of latencies and disutility). But this can not happen, since player i was content. Thus in the limit of δ approaching 0 we get $\Delta p_i/\Delta f_i \geq p_i/f_i$.

Therefore as δ tends to 0,

$$\delta + a_i \cdot \delta(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+}) = \Delta p_i \geq \frac{p_i}{f_i}\Delta f_i = \frac{p_i}{f_i} \cdot \delta(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+}),$$

$$1 + a_i(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+}) \geq \frac{p_i}{f_i}(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+}),$$

$$(\frac{p_i}{f_i} - a_i)(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v^+}) \leq 1.$$

■

We can now analyze the best response function $\beta : \mathbb{R}^k \mapsto \mathbb{R}^k$. This function maps a price vector p to another price vector p', such that p'_i maximizes player i's profit assuming all other players price as in p. We define player i's best response to be $p_i = 0$ if there is no price at which i can derive a positive profit. We will use Lemma 2.2 to show that a player's best response is unique, and hence the $\beta(p)$ is well defined. This and the continuity of $\beta(\cdot)$, which will be shown in Lemma 2.9, will be sufficient to prove the existence of pure equilibria.

Lemma 2.8 *For any set of existing prices p, all players' best response is unique (and hence the $\beta(p)$ is well defined).*

Proof : Suppose that the best response for some player i is not unique, and let $p'_i > p''_i$ be two best response prices for player i. Let f' and f'' be the corresponding flow vectors when player i selects prices p'_i and p''_i respectively, with corresponding total flow F' and F''. Lemma 2.3 implies that $f'_i < f''_i$ and $F' < F''$. Let v' denote the right slope of $u(\cdot)$ at F' and v'' denote left slope of $u(\cdot)$ at F''. By concavity of $u(\cdot)$, $v' \geq v''$ (recall that both values are negative).

Since both p_i and p''_i are best responses for player i, we can apply Lemma 2.2:

$$1 \geq (\frac{p'_i}{f'_i} - a_i)(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v'}) > (\frac{p''_i}{f''_i} - a_i)(\sum_{j\neq i} \frac{1}{a_j} - \frac{1}{v''}) \geq 1.$$

[1] To be formal, we have to take a sequence of δ's approaching 0.

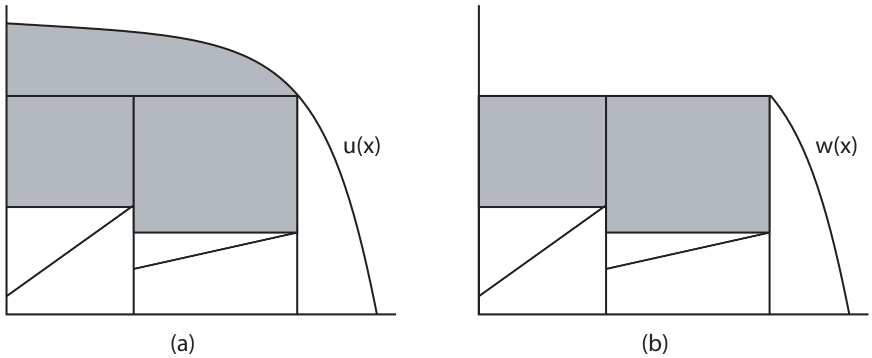

Figure 3: A disutility curve with prices and flows at equilibrium and the corresponding truncated curve.

This is a contradiction, and thus a players best response must be unique. ■

Lemma 2.9 *The best response function $\beta(\cdot)$ continuously depends on the price vector.*

Proof : It is sufficient to argue that $\beta(\cdot)$ is continuous for each player in each coordinate of the price vector. Assume not. Let $p = (p_1, \ldots, p_n)$ be a price vector for which player j's best response changes discontinuously when player i increases her price by an arbitrary small amount ϵ. Let $p' = (p_1, \ldots, p_{i-1}, p_i + \epsilon, p_{i+1}, \ldots, p_n)$ denote the new price vector, and $\beta_j(p)$ and $\beta_j(p')$ be respectively the best response prices of player j under the two price vectors. By Lemma 2.8, the best response is unique, so the profit that j gets from charging $\beta_j(p)$ under price vector p is strictly greater (say by Δ) than the profit that it would get by charging $\beta_j(p')$ instead. But we know from Lemma 2.7 that profit changes continuously in prices. Hence for a small enough ϵ, the profit difference could not reach Δ, contradicting the discontinuity assumption. ■

Since there is a maximum price P such that any player charging above P gathers no profit, we can restrict our attention on $\beta(\cdot)$ to the convex and compact region $[0, P]^k$. Thus we can apply Brouwer's fixed point theorem and thereby prove the following.

Theorem 2.10 *The network pricing game has a Nash equilibrium.*

3 Price of Anarchy

In this section we prove our main result, namely that selfish behavior on the part of the players yields a social utility that is within a small constant factor of optimal. More formally, we prove

Theorem 3.1 *The price of anarchy for the network pricing game is at most 4.65.*

The first step is to take a general instance of a game with prices at Nash equilibrium, and create a new game instance in which the equilibrium gathers no traffic utility (see Figure 3). In doing this we preserve the original equilibrium, and only increase the price of anarchy. Now, in this modified game, we can bound the social utility of an optimal solution solely against the player profit at Nash equilibria.

We will consider a general instance of the game with disutility function u and latencies $\ell_i(\cdot)$ for $1 \le i \le k$. We assume we have a price vector p at Nash equilibrium, with induced flow vector f, total flow volume F and disutility d.

Lemma 3.2 *Define a new disutility curve*

$$w(x) = \begin{cases} d & \textit{if } x < F \\ u(x) & \textit{otherwise} \end{cases}$$

Then prices p are also at Nash equilibrium given this truncated disutility curve, and the price of anarchy of this instance has not decreased.

Proof : Clearly no player has an incentive to decrease her price, as this would increase the total flow, and thus yield a flow vector that was achievable under $u(x)$. But we assumed that p was at equilibrium for $u(x)$, so this can not benefit any player. Furthermore, no player has in incentive to increase their price, as the potential gain of any such move is strictly smaller than it would have been prior to the truncation.

As for the price of anarchy, note that this truncation destroys all traffic utility of the given equilibrium, as shown in Figure 3. But this is also an upper bound on the decrease in the social utility of the new optimal solution under the truncated disutility curve. Thus the price of anarchy can only increase. ■

The next lemma provides a simple lower bound on the price that players charge at equilibrium, both in terms of the system disutility d and their own latency function. This result will clearly be useful in lower bounding player profits in the Nash equilibrium.

Lemma 3.3 *At equilibrium, any player i charges $p_i \geq \frac{d-b_i}{2}$, where $b_i = \ell_i(0)$.*

Proof : For simplicity, we will assume we have a truncated disutility curve, although the lemma is also true without this assumption. Define $q = d - b_i$. We will claim that if player i charged $p_i < \frac{q}{2}$, then she can increase her profit by charging $\frac{q}{2}$. Observe that $a_i \cdot f_i + b_i + p_i = d$. Thus we can express

$$\pi_i = p_i \cdot f_i = p_i(q - p_i)/a_i < \frac{q^2}{4a_i}.$$

However, since we are assuming a truncated disutility curve, in charging $\frac{q}{2}$, the resulting flow f' would be determined solely by $\ell_i(x)$. Hence, f' satisfies the same condition, i.e. $a_i \cdot f'_i + b_i + \frac{q}{2} = d$. Thus the profit would be exactly $\frac{q^2}{4a_i}$. ■

We now prove our main result. We consider an instance of our game with equilibrium prices p, and corresponding flow vector f, total flow F, and disutility d. We assume that our disutility function $u(x)$ is truncated as in Lemma 3.2, i.e. $u(x) = d$ on the interval $[0, F]$, and thus $U(p)$ is represented entirely by player profit.

Proof of Theorem 3.1 : As observed, in our truncated instance of the game, the given Nash equilibrium generates no traffic utility. Thus we must bound both the player profit and traffic utility of the optimal solution against the profit of the players at Nash equilibrium. We will consider an optimal price vector p^*, with corresponding flows f^*, F^*, and disutility value d^*. We can trivially bound the traffic utility of the optimal solution by $F^* \cdot (d - d^*)$, and we can attribute $f_i^* \cdot (d - d^*)$ of this bound to each player i. Thus we will think of each player i in the optimal solution as contributing $\pi_i^* + f_i^* \cdot (d - d^*)$ to $U(p^*)$.

We first partition the players by the slope of their latencies. More precisely, let us call a player *steep* if $a_i \geq \frac{1}{2}\frac{p_i}{f_i}$. Otherwise, we will say a player is *shallow*. We will now argue that there is at most one shallow player. Then we will show that the contribution of a steep player i to the optimal solution can be bounded in terms of π_i. Finally, we consider two cases regarding the shallow player's latency. In both cases, we show that the contribution of this shallow player to the optimal solution can be bounded within a constant factor of the sum of the profits of all the players at Nash equilibrium.

Intuitively, the reason why there is only a single shallow player is that if there were two (or more), then one of them would have an incentive to reduce her price slightly, since the amount of extra flow she gains by stealing from the other shallow player(s) would far outweight the slight decrease in her price. More

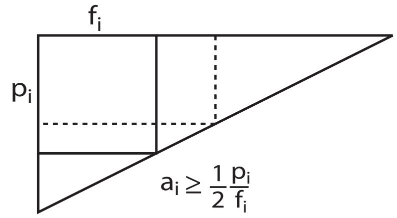

Figure 4: Bounding the value that optimum gathers from steep players.

formally, let v^+ be the right slope of $u(x)$ at F. Consider any shallow player i. By Lemma 2.2, we know that $(\frac{p_i}{f_i} - a_i)(\sum_{j \neq i} \frac{1}{a_j} - \frac{1}{v^+}) \leq 1$. Since i is shallow, the first term is greater than $\frac{p_i}{2f_i}$. But v^+ is negative, so for all $j \neq i$ we must have $a_j > \frac{p_i}{2f_i} > a_i$, as otherwise the above inequality would be violated. This implies that i must have the unique minimum slope, and thus i must be the only shallow player.

Consider a steep player j. The optimal solution can not gather more value than π_j by charging more than p_j, as then the player could selfishly do the same. Thus we only need to consider how the optimal solution might benefit by having $p_j^* < p_j$. By assumption, $a_j \geq \frac{1}{2}\frac{p_j}{f_j}$. The maximum feasible rectangle (corresponding to gathered utility) that is bounded above by $u(x)$ and below by $\ell_j(x)$ has an area of $1.125\pi_j$, which is achieved when $a_j = \frac{1}{2}\frac{p_j}{f_j}$ by setting a price of $\frac{3}{4}p_j$ and inducing a flow of $\frac{3}{2}f_j$ (see Figure 4). Thus nearly all players gather almost as much utility as they would in an optimal solution.

Now we are left with the task of bounding the value gathered by the shallow player i in the optimal solution. We again know that to gather more value with i, the optimal solution must charge a lower price and carry more flow, as any benefit in raising the price could also be realized by the player at equilibrium. Thus our concern is that somehow the optimal solution sends a huge amount of flow at lower price on i, thereby gathering substantially larger social utility. We consider two cases, based on just how shallow i's latency is.

Case 1: $\ell_i(4f_i) - \ell_i(f_i) \geq p_i/4$. In this case the latency is not very shallow, and we can simply bound the maximum contribution of player i to the optimal solution by ignoring all other players and assuming $u(x) = d$. Then the best choice for i is to charge $\frac{d-b_i}{2}$ (thereby maximizing the area of a rectangle inscribed in a triangle). The maximum area of such a rectangle would be when $\ell_i(4f_i) - \ell_i(f_i) = p_i/4$, in which case $\frac{d-b_i}{2} = \frac{13}{24}p_i$ and the corresponding flow is $\frac{13}{2}f_i$ for a total profit of $3.52\pi_i$.

Case 2: $\ell_i(4f_i) - \ell_i(f_i) < p_i/4$. This case deals with a very shallow latency. A trivial upper bound on the amount that the optimal solution can gather through i is

$$\int_0^r (u(x) - \ell_i(f_i))dx$$

where r is defined by $u(r) = \ell_i(f_i)$. This is the value i could gather in the absence of other players if $\ell_i(x)$ never exceeded $\ell_i(f_i)$. We will partition this area into two regions; A, representing $\int_F^r (u(x) - \ell_i(f_i))dx$, and B, representing the rest, with area $F(d - \ell_i(f_i))$, as shown in Figure 5(a).

To bound A, recall that $a_i < \frac{1}{2}\frac{p_i}{f_i}$. Since the slope of i's latency is shallow, the disutility curve must be relatively steep, as otherwise, i could decrease p_i slightly and dramatically increase f_i. More precisely, Lemma 2.2 implies that $v^+ \leq -\frac{1}{2}\frac{p_i}{f_i}$. Since $u(x)$ is concave, the area in A can be upper bounded by a triangle of height p_i and slope v^+. This has area at most $p_i \cdot f_i = \pi_i$.

To bound B, we partition the steep players (all $j \neq i$) into two classes. Let S^- be the set of all players j for whom $b_j > \ell_i(f_i) + p_i/2$, and let S^+ consist of all remaining steep players, as shown in Figure 5(b). Define F^- and F^+ to be the total flow carried by all players in S^- and S^+ respectively at equilibrium. By

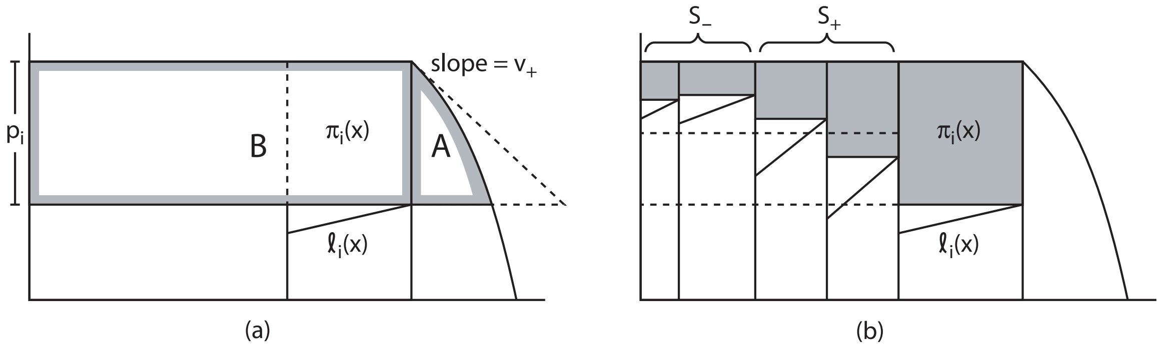

Figure 5: (a) The regions A and B for the shallow player, (b) Partitioning the steep players into S^- and S^+.

Lemma 3.3, we know that for any player $j \in S^+$, $p_j \geq (d - b_j)/2$, and hence the total profit of all players in S^+ is at least $p_i F^+/4$. Furthermore, we can argue that S^- must be small, as if it was very large, player i would have an incentive to undercut all the players in S^-. In particular, we claim that $F^- \leq 3f_i$. Otherwise, player i would have an incentive to charge a quarter of her current price. Due to the above shallowness condition on her latency, she is guarranteed to more than quadruple her flow before any player in S^- routes any traffic. Clearly this would generate more profit, contradicting our assumption of equilibrium. Thus we can bound the total area of B by $4\pi_i + 4\sum_{j \neq i} \pi_j$.

The combined area of A and B is thus at most $5\pi_i + 4\sum_{j \neq i} \pi_j$, and hence the optimal solution can gather a value of no more than $5\pi_i + 5.125\sum_{j \neq i} \pi_j$. Thus the price of anarchy is at most 5.125.

The promised bound of 4.65 is obtained by rebalancing the different bounds obtained for the different cases in the above proof. First, observe that when the slope of the shallow player was not very shallow (*Case 1* of the above proof), the bound we obtained was 3.52, which is significantly smaller 5.125. Furthermore, returning to our bound on the area of B, note that we can generalize our definition of S^- to be the set of all players j for whom $b_j > \ell_i(f_i) + \alpha p_i$ for some real α, and define S^+ similarly. Parametrizing over the slope of the shallow player a used to distinguish between *Case 1* and *Case 2*, as well as the α, we can try to minimize the maximum of the bounds obtained for the different cases. In particular, if we set $a = 0.06$ and $\alpha = 0.43$, we get the desired bound of 4.65 for the price of anarchy[2]. ■

For the special case when $\ell_i(0) = 0$ for all i, a proof similar to the one above yields the following

Theorem 3.4 *The price of anarchy of the network pricing game with a concave disutility curve and linear latencies of the form $\ell_i(x) = a_i x$ is bounded by 3.125.*

Proof : The argument will mostly follow along the lines of the proof of Theorem 3.1. As before, there will be at most a single shallow player i and we will bound the value gathered by the optimal solution from the steep players by $1.125\sum_{j \neq i} \pi_j$. As for the shallow player, we will again split her potential profit into areas A and B. Again, the area of A is at most π_i, but unlike in the proof of Theorem 3.1, we can give a simpler bound for the area B. Indeed, Lemma 3.3 implies that the area of B is at most $\pi_i + 2\sum_{j \neq i} \pi_j$, since each of the steep players charges a price of at least $\frac{d}{2}$. Hence the total cost is at most $3.125\sum_{j \neq i} \pi_j + 2\pi_i$ giving the desired guarantee. ■

[2]We also need to observe that instead of area A, we can use the smaller area of $\int_F^r (u(x) - \ell_i(f_i + x - F))dx$ to bound the utility that optimum can gather from the shallow player i. Identical to the argument in the proof of Observation 3.6, this area is at most $\frac{\pi_i}{2}$, instead of the previous bound of π_i for the area of A.

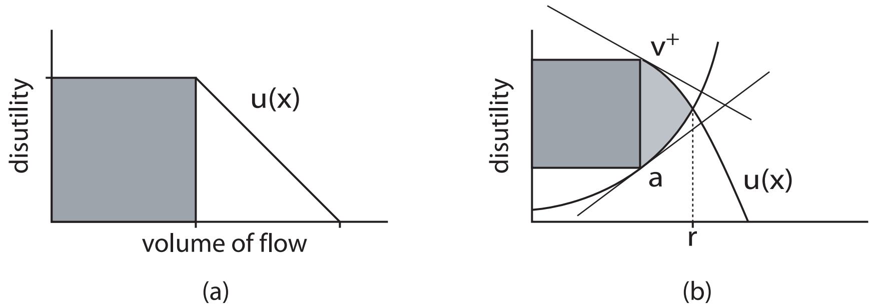

Figure 6: (a) A single player example with a price of anarchy of 3/2, (b) Proving the 3/2 upper bound on the price of anarchy of the network pricing game with a single player.

We conclude the section by presenting a single player instance of our game where the price of anarchy is 3/2. We then show that this bound it tight when there is only a single player by demonstrating a matching upper bound.

Observation 3.5 *There exists a 1-player instance of the network pricing game with linear latencies and concave disutility curve which has a price of anarchy of 3/2.*

Proof : Consider the disutility curve $u(x) = 1$ for $0 \le x \le 1$ and $u(x) = 2 - x$ for $1 \le x \le 2$ and let the player have zero latency (see Figure 6(a)). Then she would obtain maximal profit of 1 by charging a price of 1 for a social value of 1. Yet the optimal solution can gather a social value of 3/2 by charging 0. Hence the lower bound. Finally, it is not difficult to slightly modify this example to ensure strict monotonicity of latency curves. ■

Observation 3.6 *The network pricing game with a single player has a price of anarchy of at most 3/2.*

Proof : As in the proof of the main theorem, we shall assume truncated disutility curves. We will also assume without loss of generality that $p = f = 1$ for the single player (we can achieve this by rescaling the axes). As before, the only way that the optimal solution can obtain more social value than the equilibrium is by charging a lower price. Hence a trivial upper bound on the social value of the optimum is $1 + \int_1^r (u(x) - \ell(x))dx$, where r is the intersection point of $u(\cdot)$ and $\ell(\cdot)$ (see Figure 6(b)). The area of the integral can be bounded by the triangular area bounded by the slopes a and v^+ of the latency and disutility curves. Lemma 2.2 tells us that $a - v^+ \le 1$ (recall that v^+ was negative), so the area of the triangle is at most 1/2. ■

4 Extensions and Related Models

In this section we analyze what happens to the network pricing game when we relax the assumptions on the latency functions and the disutility curve. First, we consider convex, as opposed to linear, latency functions, while retaining concave disutility.

Unfortunately, the network pricing game with convex latencies and concave disutility curve may not have a Nash equilibrium as illustrated by the following 2-player example. Let the disutility curve be $u(x) = 1$. Define the player latencies as follows.

$$\ell_1(x) = \begin{cases} 0 & 0 \le x < 1 \\ \infty & x \ge 1 \end{cases} \qquad \ell_2(x) = \begin{cases} 0 & 0 \le x < \frac{1}{3} \\ \infty & x \ge \frac{1}{3} \end{cases}$$

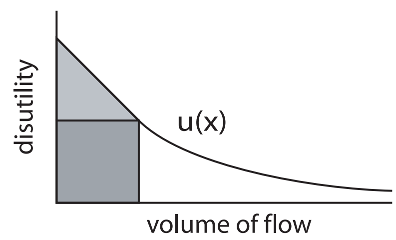

Figure 7: An example with convex disutility and unbounded price of anarchy.

We claim that this instance of the game has no Nash equilibrium. Indeed, at an equilibrium the first player can not obtain profit from the first third of the flow, since otherwise the second player would undercut him. Hence first player's best option must be to charge a price of 1 for a profit of $\frac{2}{3}$. Yet this would induce the second player to charge a high price as well, which in turn would create an incentive for the first player to undercut the second one slightly and obtain a larger profit. Although the above example violates continuity and strict monotonicity of the latency functions and the disutility curve, it is not difficult to alter it slightly so as to satisfy these conditions while still maintaining the nonexistence of Nash equilibrium.

On the other hand, when such a network pricing game does have a Nash equilibrium, the proofs of the previous sections can be extended to yield the same bounds on the price of anarchy.

Theorem 4.1 *If an instance of the network pricing game with convex latencies and concave disutility has a Nash equilibrium, then the price of anarchy is bounded by 4.65. Furthermore, if the delay is exclusively due to congestion (i.e. all links have $\ell_i(0) = 0$), then the bound can be improved to 3.125.*

We next consider instances of the game with convex, instead of concave, disutility curves, while maintaining convex latencies. An example similar to the one above can be constructed to show that this game may still not have a Nash equilibrium. Unfortunately, unlike the game with concave disutility curves, even when Nash equilibria do exist, we do not have any such bounds, since as the following single player example illustrates, there exist instances of the game with unbounded gap between the social utility of a Nash equilibrium and an optimal solution.

Example : Consider an instance of the game with a single player who has zero latency, and a disutility curve $u(x) = 2 - x$ when $0 \leq x \leq 1$ and $u(x) = \frac{1}{x}$ otherwise (see Figure 7). The equilibrium strategy for the player is to charge any price above 1 for a profit of 1 and finite social utility, yet charging a price of 0 yields and infinite social benefit. ■

References

[1] D. Acemoglu and A. Ozdaglar. Competition and Efficiency in Congested Markets. *Unpublished Manuscript*, 2004.

[2] D. Acemoglu and A. Ozdaglar. Flow Control, Routing, and Performance from Service Provider Viewpoint. *Unpublished Manuscript*, 2004.

[3] P. Afèche. Incentive-Compatible Revenue Management in Queuing Systems: Optimal Strategic Idleness and other Delaying Tactics. *Unpublished Manuscript*, 2004.

[4] G. Allon and A. Federgruen. Competition in Service Industries. *Unpublished Manuscript*, 2004.

[5] E. Anshelevich, A. Dasgupta, J. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. *IEEE Symposium on Foundations of Computer Science*, 295–304, 2004.

[6] E. Anshelevich, A. Dasgupta, É. Tardos, and T. Wexler. Near-Optimal Network Design with Selfish Agents. *ACM Symposium on Theory of Computing*, 511–520, 2003.

[7] M. Beckmann, C. B. McGuire, and C. B. Winsten. Studies in the Economics of Transportation. *Yale University Press*, 1956.

[8] R. Cole, Y. Dodis, and T. Roughgarden. Pricing Network Edges for Heterogeneous Selfish Users. *ACM Symposium on Theory of Computing*, 521–530, 2003.

[9] R. Cole, Y. Dodis, and T. Roughgarden. How Much Can Taxes Help Selfish Routing. *ACM Conference on Electronic Commerce*, 98–107, 2003.

[10] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish Routing in Capacitated Networks. *Mathematics of Operations Research*, 961–976, 2004.

[11] A. Czumaj, P. Krysta and B. Vöcking. Selfish Traffic Allocation for Server Farms. *ACM Symposium on Theory of Computing*, 287–296, 2002.

[12] A. Czumaj and B. Vöcking. Tight Bounds for Worst-Case Equilibria. *ACM-SIAM Symposium on Discrete Algorithms*, 413–420, 2002.

[13] A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker. On a Network Creation Game. *ACM SIGACT-SIGOPT Symposium on Principles of Distributed Computing*, 247–251, 2003.

[14] L. Fleischer, K. Jain, and M. Mahdian. Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games. *IEEE Symposium on Foundations of Computer Science*, 277–285, 2004.

[15] G. Karakostas and S. G. Kolliopoulos. Edge Pricing of Multicommodity Networks for Heterogeneous Selfish Users. *IEEE Symposium on Foundations of Computer Science*, 268–276, 2004.

[16] E. Koutsoupias and C. H. Papadimitriou. Worst-Case Equilibria. *International Symposium on Theoretical Aspects of Computer Science*, 404–413, 1999.

[17] I.H. Lee and R. Mason. Market Structure in Congestible Markets. *European Economic Review*, 45 (4-6), 809–818, 2001.

[18] A. C. Pigou. The Economics of Welfare. *Macmillan and Co.*, 1920.

[19] T. Roughgarden. The Price of Anarchy is Independent of the Network Topology. *ACM Symposium on Theory of Computing*, 428–437, 2002.

[20] T. Roughgarden. Stackelberg Scheduling Strategies. *ACM Symposium on Theory of Computing*, 104–113, 2001.

[21] T. Roughgarden and É. Tardos. How Bad Is Selfish Routing? *IEEE Symposium on Foundations of Computer Science*, 93–102, 2000.

[22] S. Suri, C. D. Toth, and Y. Zhou. Selfish Load Balancing and Atomic Congestion Games. *ACM Symposium on Parallel Algorithms and Architectures*, 188–195, 2004.

[23] A. Vetta. Nash Equilibria in Competitive Societies, with Applications to Facility Location, Traffic Routing and Auctions. *IEEE Symposium on Foundations of Computer Science*, 416–425, 2002.